μCeiver-Fi: Exploiting Spectrum Resources of Multi-link Receiver for Fine-Granularity Wi-Fi Sensing

Xin Li Yinghui He* Jun Luo
College of Computing and Data Science, Nanyang Technological University (NTU), Singapore
Email: {l.xin, yinghui.he, junluo}@ntu.edu.sg

ABSTRACT

Wi-Fi is deemed as a promising sensing media due to its ubiquity, yet Wi-Fi sensing is known to be confined by its limited bandwidth that leads to insufficient range resolution. Though sampling a wider spectrum multiple times can enable wideband sensing, its practicality is still hampered by the need for accessing Wi-Fi firmware. In this paper, we propose μCeiver-Fi to exploit spectrum resources for fine-granularity Wi-Fi sensing; it relies solely on a commodity *multi-link receiver*. Since the channel samples from multiple links under the same receiver can still be misaligned, we first innovate in a comprehensive calibration process to align these samples. This is followed by a novel optimization framework to extend effective sensing bandwidth to GHz-level using only a few channel samples. Finally, we specifically design a spectral representation for sensing information in order to bridge between wideband signals and diversified downstream applications. Through comprehensive evaluations in Wi-Fi pose estimation task, we demonstrate the promising performance of μCeiver-Fi in fine-granularity sensing.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing**; • **Hardware** → **Signal processing systems**.

KEYWORDS

Wi-Fi human sensing, Wi-Fi human pose estimation, multi-link operation, range resolution, ISAC.

ACM Reference Format:
X. Li, Y. He, and J. Luo. 2025. μCeiver-Fi: Exploiting Spectrum Resources of Multi-link Receiver for Fine-Granularity Wi-Fi Sensing. In *The 31st Annual International Conference on Mobile Computing and Networking (ACM MobiCom'25), November 4–8, 2025, Hong Kong, China.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3680207.3765263

1 INTRODUCTION

Wi-Fi has undergone significant advancements over the past decade [8, 16, 18, 25, 61, 71, 73] and has gradually become an indispensable part of modern life [65]. The ubiquitous presence of Wi-Fi signals drives researchers to conduct comprehensive examinations and studies [28, 68, 71], leveraging it as more than just a communication tool to fully harness this resource. Across various research directions, pervasive sensing is widely regarded as a fundamental capability that Wi-Fi is destined to acquire [34, 37], thanks to the availability of channel state information (CSI) [15, 23]. As outstanding applications, localization [6, 9, 51, 76], tracking [45, 46, 54, 67], human activity recognition [12, 21, 58, 75], vital sign detection [17, 35, 62, 72], and imaging [22, 52, 59, 63] have all demonstrated promising results, showcasing the practical significance of Wi-Fi sensing.

However, Wi-Fi sensing is inherently constrained by the limited bandwidth, preventing it from achieving sufficient resolution for fine-granularity sensing [2], as illustrated in Figure 1 (top panel). Essentially, CSI samples acquired within a narrow bandwidth offer very limited *diversity*, so they may fail to capture subtle multipath channel variations necessary to distinguish close subjects and/or to detect slight movements [9]. While narrow bandwidth severely impacts sensing

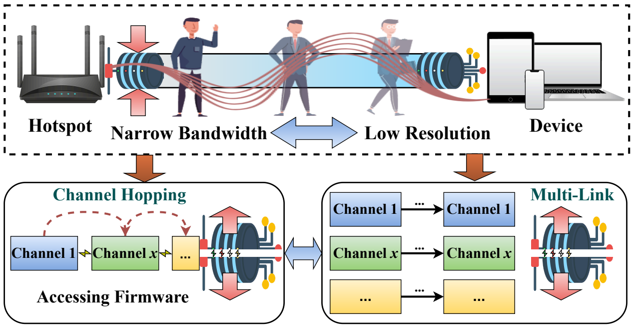


Figure 1: Channel hopping requires firmware access, restricting its realization on commodity devices. Fortunately, multi-channel samples can be simultaneously collected by commodity multi-link receivers.

---

* Yinghui He is the corresponding author.

ACM MobiCom'25, November 4–8, 2025, Hong Kong, China

ACM ISBN 979-8-4007-1129-9/2025/11
https://doi.org/10.1145/3680207.3765263

performance, endowing excessively wide bandwidth to Wi-Fi can be unrealistic [53], as its ubiquity forces standards to restrain bandwidth in order to address co-channel interference concerns for the primary communication purpose [10]. Although multiple antennas [22, 24, 72] or devices [18, 47, 54] can partially alleviate the diversity limitation caused by narrowband, the incurred high cost and cumbersome in deployment can always hinder their wide adoption.

Compared with other makeshifts, fully exploiting available spectrum resources is still the most effective solution [30, 57, 66, 69]. Since early methods [57, 66, 69] are confined by channel availability and efficiency, the latest proposals [29, 30] adopt a fast channel hopping scheme to collect discrete channel samples and integrates them into a spectrum with impressive GHz-level sensing bandwidth, as shown in Figure 1 (lower-left panel). However, the channel hopping scheme requires firmware access, yet the increasingly closed nature of driver designs complicates its adaptation to up-to-date network interface cards (NICs). Moreover, sequential collection of multiple channel samples limits the number of observations achievable within channel coherence time, reducing its applicability in unstable environments. Given these limitations, employing multi-link receivers to simultaneously collect multiple channel samples appears to be a more plausible approach, as demonstrated by a recent proposal [43]. Unfortunately, this prototype, implemented solely on millimeter-wave software-defined radio (SDR), does not account for complex interference from commodity Wi-Fi chips, leaving the feasibility of commodity multi-link receivers unexplored.

In practice, the latest commodity Wi-Fi NICs [1, 3] are capable of connecting to multiple channels simultaneously, providing a foundation for expanding bandwidth free of firmware access, as shown in Figure 1 (right panel). However, this intuitively attractive approach faces its own challenges. Firstly, channel samples from the multi-link receiver remain misaligned, as they experience different interference [5], making it impossible to directly integrate them into a wideband signal. Secondly, even with successful integration, the resulting bandwidth falls short of the GHz-level required for fine-granularity sensing, as current commodity NICs are typically restricted to 2 simultaneous channel accesses across different bands (e.g., 2.4 GHz and 5 GHz) [3]. Finally, since wideband signals are employed by various general sensing tasks, it is essential to provide general representation to enable adaptation across diverse applications.

To address the above challenges, we propose μCeiver-Fi to leverage commodity multi-link receivers for expanding the effective bandwidth of Wi-Fi sensing. We first categorize the interference in CSIs into time-varying, time-invariant, and residual components and design disentanglement and calibration algorithms to fully eliminate them for channel sample alignment. Building on this, we develop an optimization model to extract frequency-independent features from clean multi-link CSIs and use them to reconstruct the wideband signal. To support various downstream sensing applications, we design a spectral representation to capture subject state information; it enables a neural network to readily bridge the wideband signal with application-oriented data structure. Finally, we implement a prototype of μCeiver-Fi with the latest commodity NICs, capable of simultaneously collecting channel samples on the 2.4 GHz and 5 GHz bands, and conduct comprehensive experiments to evaluate its performance on the promising yet challenging Wi-Fi human pose estimation task. Our main contributions are as follows:

- We propose μCeiver-Fi system to achieve fine-granularity Wi-Fi sensing with the commodity multi-link receiver.
- We develop a progressive framework to gradually disentangle interference across multiple channel samples for precise alignment.
- We build an optimization model to reconstruct wideband signals from two available channel samples.
- We design a method to embed subject state information into the spectrum, providing a general representation for diverse sensing applications.
- We implement a prototype of μCeiver-Fi and conduct comprehensive evaluations to demonstrate its sensing capabilities using the human pose estimation task.

Our μCeiver-Fi expands the effective bandwidth to realistically attain sufficient range resolution, rather than recovering the complete channel from 2.4 GHz to 5 GHz. In other words, the information present in the unseen frequencies between these bands may not be fully or accurately captured.

The rest of our paper is structured as follows. Section 2 introduces the basics Wi-Fi sensing model and the motivation behind μCeiver-Fi. Section 3 presents the system design of μCeiver-Fi. Section 4 details the prototype implementation and experiment setup of μCeiver-Fi. Section 5 reports the evaluation results. Related work and discussions are outlined in Section 6. The conclusion of our paper is in Section 7.

2 PRELIMINARY AND MOTIVATION

In this section, we first establish the basic Wi-Fi sensing model and explain the key factors concerning fine-granularity sensing. We then analyze the limitations of channel hopping and highlight the potential of multi-link receivers for bandwidth expansion. Finally, we discuss the challenges of achieving fine-granularity sensing with multi-link channel samples on commodity Wi-Fi NICs.

2.1 Wi-Fi Sensing Preliminary

Assuming a Wi-Fi system operating in a multipath scenario, the received signal at the receiver (Rx) can be characterized by a set of parameters $(\tau_k, \theta_k)_{k\in[1,\,K]}$, where τ_k and θ_k represent the time of flight (ToF) and angle of arrival (AoA) of the k-th path, respectively. For a system with M subcarriers and an Rx equipped with N antennas arranged in a uniform array, the CSI $\boldsymbol{H}^{\text{ideal}} \in \mathbb{C}^{N \times M}$ can be modeled as [30]:

$$\begin{aligned} h^{\text{ideal}}_{n,m} &= \textstyle\sum_{k=1}^{K} \alpha_{n,m,k} \cdot h^{\tau}_{m,k} \cdot h^{\theta}_{n,k} \\ &= \textstyle\sum_{k=1}^{K} \alpha_{n,m,k} \cdot \mathrm{e}^{-\mathrm{i}2\pi(f^{\mathrm{c}} \pm (m-1) f^{\mathrm{b}})\tau_k} \cdot \mathrm{e}^{-\mathrm{i}2\pi(n-1) d \cos(\theta_k) f^{\mathrm{c}}/\mathrm{c}}, \end{aligned} \tag{1}$$

where α denotes amplitude attenuation, n and m are the antenna and subcarrier indices, respectively, f^{c} and f^{b} are the carrier frequency and subcarrier bandwidth, respectively, d is the distance between neighboring antennas, and c denotes the speed of light. This model indicates that CSI $\boldsymbol{H}^{\text{ideal}}$ fuses both line-of-sight (LoS) and non-line-of-sight (nLoS) paths into one matrix. In order to identify individual paths (a.k.a. fine *range resolution*) out of $\boldsymbol{H}^{\text{ideal}}$, one needs to have a fine *temporal resolution* $\Delta\tau$ in measuring τ_k, and $\Delta\tau = 1/B$ improves with bandwidth B [2], highlighting the importance of bandwidth expansion for fine-granularity Wi-Fi sensing[1].

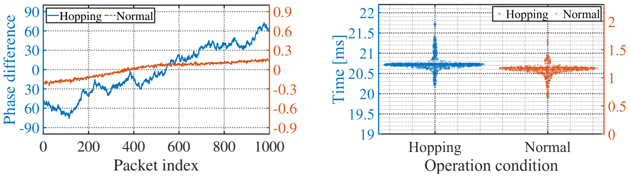
Meanwhile, in real-world scenarios, signals received by commodity Wi-Fi NICs not only contain noise but are also affected by various hardware-induced offsets and boundary effects [7]. Thus, the real-world CSI $\boldsymbol{H}$ becomes:

$$h_{n,m} = \beta_m \cdot h^{\text{ideal}}_{n,m} \cdot \mathrm{e}^{-\mathrm{i}2\pi\left(\phi^{\mathrm{D}}_{t,n,m} + \phi^{\mathrm{S}}_{n,m}\right)} + \epsilon_t, \tag{2}$$

where t denotes the index of data packet (*slow-time*); $\phi^{\mathrm{D}}_{t,n,m}$ and $\phi^{\mathrm{S}}_{n,m}$ represent *time-varying* offsets with respect to t and *time-invariant* offsets with respect to subcarriers and antennas, respectively. In general, the environment noise ϵ is independent and identically distributed across packets. Moreover, the boundary effect β_m alters the subcarriers at both ends of the CSI. Such complex interferences significantly affect fine-granularity sensing, thus making bandwidth expansion from narrow channels extremely challenging.

2.2 Channel Hopping vs. Multi-Link

A recent proposal UWB-Fi [30] employs sparse channel hopping to sample CSIs across a 4.7 GHz bandwidth. To enhance the efficiency of channel hopping, UWB-Fi requires NIC kernel access; yet, as chip designs become increasingly closed [20], implementing UWB-Fi on future NICs would need unwarranted hacking. Moreover, channel hopping requires restarting the Wi-Fi NIC, introducing random initial phase offsets that result in significantly higher interference compared to normal operation. To demonstrate this, we collect samples during channel hopping using the method described in [30] and also perform normal sampling within a particular channel in the same environment. Figure 2(a) shows the unwrapped phase difference between signals from adjacent antennas for both methods: the phase difference under normal sampling remains several hundred times smaller than that under channel hopping. Due to these random phase offsets, the resolution achieved by $\Delta\tau$ via a wide bandwidth B can be completely annihilated.

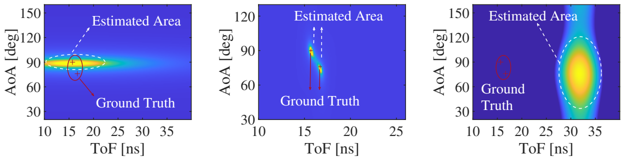


(a) Introduced offsets. (b) Sampling time.

Figure 2: Channel hopping vs. normal sampling in terms of (a) introduced offsets and (b) sampling time.

In addition, due to its sequential sampling strategy, a complete observation with the channel hopping requires at least 20 ms to ensure accurate sensing in typical indoor scenarios [30], whereas capturing a single channel sample takes only 1 ms, as shown in Figure 2(b). This confines the number of complete observations within channel coherence time, posing challenges in highly unstable environments (also see Section 5.3.5). Therefore, it is desirable to sample multiple channels simultaneously using the same Rx of a commodity Wi-Fi NIC, for the sake of both effectiveness and efficiency.

Fortunately, the latest multi-link receivers support simultaneous connections to multiple links across different bands [1, 3], so they may potentially meet the aforementioned need, as partially demonstrated by HiSAC [43]. To validate its benefits, we simulate a two-subject localization scenario. We first consider SpotFi [25] with 40 MHz single-channel CSI sampled at the carrier frequency of 5230 MHz: as shown in Figure 3(a), the limited bandwidth fails to distinguish two closely spaced subjects. We then switch to two 20 MHz channels at carrier frequencies of 2472MHz and 5220MHz, and we adopt HiSAC [43] to fuse the two simultaneously sampled CSIs into a wideband observation: Figure 3(b) indicates that significantly improved resolution enables a clear distinction between the two subjects based on their peak values. These experiments corroborate what have been claimed in [43] about the benefit of exploiting multi-link receivers.



(a) Single channel. (b) Multi-link. (c) Real-world.

Figure 3: Localization results for (a) simulated single channel sample, (b) simulated multi-link channel samples, and (c) real-world multi-link channel samples.

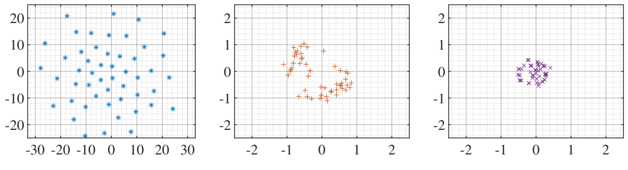
---

[1] To highlight the role of effective sensing bandwidth, we focus on static and quasi-static subjects in this work. For scenarios involving dynamic movements, Doppler frequency shifts (DFS) can serve as an orthogonal factor, which we leave for future work.

2.3 Presence of Residual Offsets

We obtain the aforementioned results via simulations because real-world multi-link channel samples from commodity Wi-Fi NICs fail to reveal meaningful subject information, as shown in Figure 3(c). The reason for this is two-fold. On one hand, HiSAC [43], validated only on mmWave-SDR, may not account for the low directivity and complex interference inherent in commodity Wi-Fi NICs. On the other hand, HiSAC's calibration method that aims to align the CSIs from multiple links seems to have left an unpredictable *residual offset* for each link, which in turn results in the substantial location bias shown in Figure 3(c). The issue with residual offsets stems from HiSAC's calibrating an estimated LoS path length (basically a localization procedure) against its known value: because Wi-Fi localization leverages only *phase differences* (i.e., ToF estimated from subcarrier phase differences and AoA from antenna phase differences) but neglects absolute phase values (exhibit randomness), certain absolute offsets shared across subcarriers and antennas are not taken into account by the calibration, thus leading to a residual offset $\mathrm{e}^{-\mathrm{i}2\pi\Delta_\ell}$ between the true CSI $\boldsymbol{H}_\ell^{\mathrm{ideal}}$ for the ℓ-th link and $\hat{\boldsymbol{H}}_\ell$ calibrated by HiSAC: $\hat{\boldsymbol{H}}_\ell = \mathrm{e}^{-\mathrm{i}2\pi\Delta_\ell}\boldsymbol{H}_\ell^{\mathrm{ideal}}$.

This residual offset certainly differs across different links due to distinct channel properties, yet its unpredictability persists even for the same link. To demonstrate this, we execute HiSAC's calibration method for 50 continuous CSI samples from the same link and visualize the resultant offset vectors for each sample (formed by scalar offsets for respective subcarriers and antennas), using t-distributed Stochastic Neighbor Embedding (t-SNE) [38] in Figure 4(a); the evident randomness of the residual offset (even for the same link) is clearly observable. This mainly stems from the insufficient constraints considered by HiSAC's calibration that fail to capture Δ_ℓ. Intuitively speaking, adding proper constraints (as we shall do) should enable the calibrated offsets, shown in Figure 4(b), to form compact clusters while remaining distinctly separated from the ideal CSI phase vectors shown in Figure 4(c), indicating consistent calibration results for the same link ℓ. Due the major differences between SDR and commodity Wi-Fi such as the additional interference described in Eqn. (2), HiSAC still cannot be applied to commodity Wi-Fi even with a proper calibration. Therefore, these challenges motivate the entirely novel algorithmic design in μCeiver-Fi to expand the effective sensing bandwidth.



(a) Calibrated offsets. (b) Constrained offsets. (c) Ideal CSI phases.

Figure 4: Visualization with t-SNE. (a) Offsets calibrated by HiSAC appear random, while (b) those calibrated with proper constrains remain consistent, and distinctly separated from (c) ideal CSI phase components.

3 SYSTEM DESIGN

Our μCeiver-Fi leverages commodity multi-link receiver to expand the effective sensing bandwidth for Wi-Fi sensing. It consists of three main components as illustrated in Figure 5:

- Interference Disentanglement: It clearly categorizes interference and disentangles them progressively, aiming to recover the $\boldsymbol{H}_\ell^{\mathrm{ideal}}$.
- Bandwidth Reconstruction: An optimization approach to extract features from multiple channel samples and reconstruct them into a wideband signal.
- Fine-granularity Sensing Adapter: A general spectral representation embedding subject state information to serve as the output of neural networks, enabling adaptability to diverse sensing tasks.

The processing pipeline starts with μCeiver-Fi receiveing channel samples from different channel frequencies using the multi-link receiver. Samples are then progressively disentangled to fully eliminate interferences that include, in particular, environment noise and hardware-induced offsets, as presented in Section 3.1. Leveraging the calibrated CSIs, an optimization approach is exploited to extract frequency-independent components and reconstruct them into CSIs with GHz-level bandwidth, as discussed in Section 3.2. Finally, a deep learning model transforms the wideband signals into spectra embedded with subject information, capable of adapting to various sensing tasks, though only a representative task is explored in Section 3.3 and later experiments.

3.1 Interference Disentanglement

To completely disentangle interference for retrieving clean CSI samples per link, we propose a three-step approach. First, we adopt the low-rank matrix [50] constraint to remove time-varying interference. We then design a search

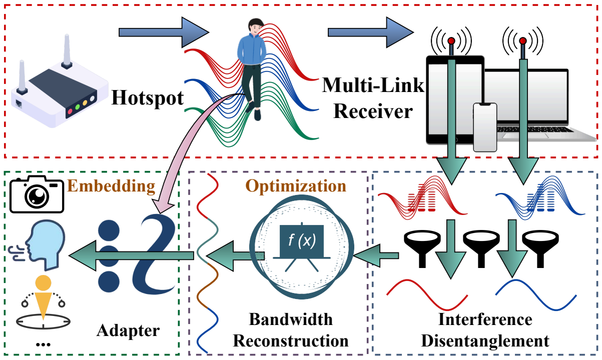


Figure 5: μCeiver-Fi system overview.

method with carefully crafted control parameters to eliminate time-invariant interference. Finally, we introduce a subcarrier random sampling scheme to generate shuffled subcarrier samples that aims to reintroduce the low-rank matrix constraint for removing the residual offsets.[2]

3.1.1 Time-varying Interference Disentanglement. Certain offsets, such as carrier frequency offset (CFO) and packet detection delay (PDD), vary over time in commodity Wi-Fi NICs. Meanwhile, the independent environmental noise and a portion of inter-subcarrier interference exhibit variations across Wi-Fi packets. Therefore, we collectively refer to them as *time-varying interference*. For ease of expression, let $\boldsymbol{H}^{\mathrm{S}} \in \mathbb{C}^{(2NM)\times T}$ denote the matrix constructed from T CSI samples $\boldsymbol{H}$ for each link, where each column corresponds to a sample and is formed by concatenating the amplitude and phase of the vectorized $\boldsymbol{H}$, and then it can be regarded as the combination of a matrix $\boldsymbol{H}^{\mathrm{S,L}}$ (containing the ideal CSIs) and a matrix $\boldsymbol{H}^{\mathrm{S,P}}$ (containing time-varying interference): $\boldsymbol{H}^{\mathrm{S}} = \boldsymbol{H}^{\mathrm{S,L}} + \boldsymbol{H}^{\mathrm{S,P}}$, according to Eqn. (2). Our method for disentangling time-varying interference leverages the constraint that ideal CSIs within the channel coherence time remain stable [55] and exhibit high correlation, which means $\boldsymbol{H}^{\mathrm{S,L}}$ is low-rank, while the variations introduced by time-varying interference are minimal and thus $\boldsymbol{H}^{\mathrm{S,P}}$ can be regarded as a perturbation [50]. Specifically, we can represent $\boldsymbol{H}^{\mathrm{S,L}}$ using an ordinary orthonormal subspace $\boldsymbol{D} \in \mathbb{C}^{(2NM)\times\Gamma}$ and a coefficient matrix $\boldsymbol{\kappa} \in \mathbb{C}^{\Gamma\times T}$: $\boldsymbol{H}^{\mathrm{S,L}} = \boldsymbol{D}\boldsymbol{\kappa}$. Thus, the recovery of $\boldsymbol{H}^{\mathrm{S,L}}$ can be formulated as a rank minimization problem, which is equivalent to minimizing the group sparsity of $\boldsymbol{\kappa}$ [50]:

$$\min_{\boldsymbol{H}^{\mathrm{S,P}},\boldsymbol{D},\boldsymbol{\kappa}} \|\boldsymbol{\kappa}\|_{\text{row-1}} + \lambda\|\boldsymbol{H}^{\mathrm{S,P}}\|_1, \text{ s.t. } \boldsymbol{D}\boldsymbol{\kappa} + \boldsymbol{H}^{\mathrm{S,P}} = \boldsymbol{H}^{\mathrm{S}}, \boldsymbol{D}^{\mathsf{T}}\boldsymbol{D} = \mathbf{I}, \quad (3)$$

where $\|\cdot\|_{\text{row-1}}$ represents row-1 norm [14] and λ is the weight parameter for the perturbation matrix.

To solve this problem, we transform it into an augmented Lagrange function:

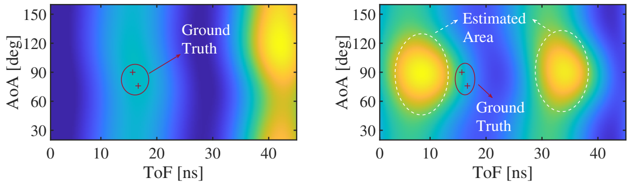
$$\begin{aligned}\mathcal{L} =& \|\boldsymbol{\kappa}\|_{\text{row-1}} + \lambda\|\boldsymbol{H}^{\mathrm{S,P}}\|_1 + \nu\left(\boldsymbol{H}^{\mathrm{S}} - \boldsymbol{D}\boldsymbol{\kappa} - \boldsymbol{H}^{\mathrm{S,P}}\right)\\ &+ \mu\|\boldsymbol{H}^{\mathrm{S}} - \boldsymbol{D}\boldsymbol{\kappa} - \boldsymbol{H}^{\mathrm{S,P}}\|_F^2/2, \quad \text{s.t. } \boldsymbol{D}^{\mathsf{T}}\boldsymbol{D} = \mathbf{I},\end{aligned} \quad (4)$$

where μ and ν are the over-regularization parameter and Lagrange multiplier, respectively, and $\|\cdot\|_F$ represents Frobenius norm [14]. Despite being a non-convex problem, Eqn. (4) can still be effectively solved, for example, using the alternating direction method, which iteratively updates $\boldsymbol{D}$, $\boldsymbol{\kappa}$, $\boldsymbol{H}^{\mathrm{S,P}}$, and ν, and exhibits strong convergence behavior in practice [50]. Moreover, its complexity scales only with the matrix dimensions and rank, thereby achieving high efficiency. After solving the problem, we remove $\boldsymbol{H}^{\mathrm{S,P}}$ from $\boldsymbol{H}^{\mathrm{S}}$ to obtain $\boldsymbol{H}^{\mathrm{S,L}} = \boldsymbol{H}^{\mathrm{S}} - \boldsymbol{H}^{\mathrm{S,P}}$, and further refine $\boldsymbol{H}^{\mathrm{S,L}}$ by repeating the above process with $\boldsymbol{H}^{\mathrm{S,L}} \rightarrow \boldsymbol{H}^{\mathrm{S}}$.

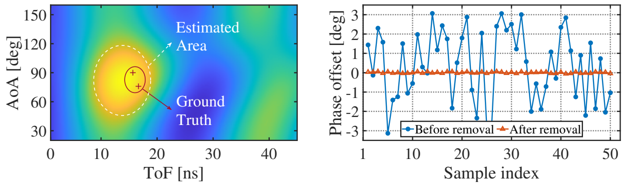
To evaluate the effectiveness of this algorithm, we collect CSI samples with a carrier frequency of 5190 MHz and a 40 MHz bandwidth in an indoor environment containing 2 subjects. Figure 6(a) presents the estimation result on raw CSIs, which fails to extract meaningful information. After disentangling for time-varying interference, Figure 6(b) reveals two distinct peaks that gradually approach the ground truth, indicating partial interference removal. Of course, due to remaining interference, further disentanglement is required for accurate estimation.

3.1.2 Removing Time-invariant Interference. The CSI still contains distortions independent of slow-time (or packets), such as sampling frequency offset (SFO), carrier phase offset (CPO), part of the inter-subcarrier interference, and boundary effects. These distortions are collectively referred to as *time-invariant interference*. Since phase variations across subcarriers and antennas are leveraged to perform LoS path estimation, HiSAC [43] uses LoS path estimation to calibrate phases (thus estimating CSIs); this, unfortunately, has been demonstrated in Section 2.3 as ineffective for commodity Wi-Fi due to the resultant residual offsets. To this end, we purposely introduce random biases into the estimated CSI while borrowing the LoS path calibration idea of HiSAC. The consequent diversity in residual offsets can then be exploited to remove them in the next step.

Since time-invariant interference affects CSI along subcarrier and antenna dimensions, we model it using two vectors $\boldsymbol{g}^{\mathrm{ant}} \in \mathbb{C}^{N}$ and $\boldsymbol{g}^{\mathrm{sub}} \in \mathbb{C}^{M}$, and each CSI sample can be represented as $\boldsymbol{H}^{\mathrm{S,L}} = \mathrm{diag}(\boldsymbol{g}^{\mathrm{ant}})\check{\boldsymbol{H}}^{\mathrm{S,L}}\mathrm{diag}(\boldsymbol{g}^{\mathrm{sub}})$, where $\boldsymbol{H}^{\mathrm{S,L}} \in \mathbb{C}^{N\times M}$ is the restructured matrix of a single CSI entry



(a) Raw CSI. (b) W/o time-varying interference.



(c) W/o time-invariant interference. (d) Residual offsets.

Figure 6: Interference disentanglement: (a)-(c) estimation results across the process and (d) residual offsets.

---

[2] In Sections 3.1.1 and 3.1.2, we perform individual interference disentanglement for each link, and in Section 3.1.3, we jointly process the CSI from the multiple links to achieve residual offsets disentanglement.

after removing time-varying interference, and $\mathbf{diag}(\cdot)$ denotes the operation of constructing a diagonal matrix from a given vector. To obtain the CSI without time-invariant interference, $\check{\boldsymbol{H}}^{\mathrm{S,L}}$, our method leverages the measured LoS path $(\tau^{\mathrm{LoS}}, \theta^{\mathrm{LoS}})$ and aims to minimize its estimation error, as

$$\varepsilon = |\theta^{\mathrm{LoS}} - \hat{\theta}^{\mathrm{LoS}}|/\theta^{\mathrm{LoS}} + |\tau^{\mathrm{LoS}} - \hat{\tau}^{\mathrm{LoS}}|/\tau^{\mathrm{LoS}}, \tag{5}$$

where $\hat{\tau}^{\mathrm{LoS}}$ and $\hat{\theta}^{\mathrm{LoS}}$ are ToF and AoA estimated from the CSI using a conventional algorithm (e.g., MUSIC [49]) under a given pair of g^{ant} and g^{sub}, respectively. Specifically, we first initialize Ψ candidate pairs, denoted as $\mathcal{U} = \{u_1, \cdots, u_\psi, \cdots\}$, where $u_\psi = \{\boldsymbol{g}_\psi^{\mathrm{ant}}, \boldsymbol{g}_\psi^{\mathrm{sub}}\}$. In the r-th iteration, we calculate the estimation error for each pair. The $\max(\Psi{-}r, \Psi/10)$ pairs with the lowest errors are reserved in the set $\mathcal{U}$ while the remaining pairs are removed. Next, we select the pair with the smallest error, denoted by $\psi^\star$, and use it to update the pairs with errors greater than c_1. For all elements in the pair pending update, the update process consists of two steps: i) the element is randomly replaced with another element from any other pair, and ii) the element is updated using the corresponding element from pair $u_{\psi^\star}$ as

$$g_\psi^{(r+1)} = \begin{cases} g_\psi^{(r)}(1-c_2)+\left(g_{\psi^\star}^{(r)}+g^{\mathrm{u}}c_3\right)c_2, & |g_\psi^{(r+1)}| \le g^{\mathrm{u}}, \\ g^{\mathrm{u}} e^{\mathrm{i}\angle g_\psi^{(r+1)}}, & \text{otherwise,} \end{cases} \tag{6}$$

where $g_\psi^{(r)}$ denotes the element in u_ψ for the r-th iteration, and g^{u} represents the upper bound. For each iteration, parameter c_2 is randomly chosen as either 0 or 1, and c_1 and c_3 are random coefficients within $[0,1]$ and $[0, 1-\sqrt{r/R}]$, respectively, with R being the maximum number of iterations. We repeat the above process until the minimum error is sufficiently small or the number of iterations reaches R, and then $\check{\boldsymbol{H}}^{\mathrm{S,L}}$ is obtained by the estimated interference vectors $\boldsymbol{g}^{\mathrm{ant}}$ and $\boldsymbol{g}^{\mathrm{sub}}$. Essentially, our method augments the search space by replacing and updating (interference) candidates to substantially improve the searching effectiveness and guarantee convergence, while its computational complexity depends on R, Ψ, N, and M [39]. Figure 6(c) presents the localization results of our method; it indicates that the peak area starts to overlap with the ground truth area, confirming the effectiveness of removing time-invariant interference.

3.1.3 Residual Offsets Disentanglement. Since the previous search method introduces randomly controlled parameters (i.e., c_1, c_2, c_3), it leaves high-diversity residual offsets in the CSI samples. These offsets apparently vary with CSI samples, so it seems reasonable to treat them as perturbations and disentangle them using the method described in Section 3.1.1. However, as each residual offset is applied to the entire CSI sample, its variation does not affect the correlation between two arbitrary CSI samples, rendering the low-rank constraint inapplicable. To this end, we randomly permute elements corresponding to the same subcarrier across CSI samples to perturb the correlations among residual offsets; this does not affect the correlations between ideal CSIs because they should exhibit high consistency by default [56]. To further enhance diversity, we also concatenate the CSIs from different links to jointly process them.

Algorithm 1: Disentanglement with Shuffled Samples

```
Input: Channel samples containing only residual
       offsets $\check{\boldsymbol{H}}^{\mathrm{S,L}}$.
Output: Clean CSI matrix $\tilde{\boldsymbol{H}}^{\mathrm{S,L}}$.
Initialization: $V = 2\sum_\ell NM_\ell$, $T$ is the sample number,
    $\boldsymbol{H}^{\mathrm{sf}} = \emptyset$, and $\boldsymbol{H}^{\mathrm{ct}} = \emptyset$.
for $v = [1, \cdots, V]$ do
    Shuffle the packet index list $[1, \cdots, T]$ randomly
      to generate $A_v$;
    $\boldsymbol{H}^{\mathrm{sf}}(v, [1, \cdots, T]) = \check{\boldsymbol{H}}_{\mathrm{L}}(v, A_v)$;
end
$\boldsymbol{H}^{\mathrm{ct}} \leftarrow \check{\boldsymbol{H}}^{\mathrm{S,L}} \cup \boldsymbol{H}^{\mathrm{sf}}$, with indices shuffled along the
    packet dimension;
Disentangle $\boldsymbol{H}^{\mathrm{ct}}$ based on Eqn. (4);
Extract samples corresponding to $\check{\boldsymbol{H}}^{\mathrm{S,L}}$ to obtain $\tilde{\boldsymbol{H}}^{\mathrm{S,L}}$.
```

Specifically, the CSIs from different links are first concatenated into a single matrix $\check{\boldsymbol{H}}^{\mathrm{S,L}} = [\check{\boldsymbol{H}}_1^{\mathrm{S,L}}; \cdots; \check{\boldsymbol{H}}_\ell^{\mathrm{S,L}}; \cdots] \in \mathbb{C}^{(2\sum_\ell NM_\ell)\times T}$, where M_ℓ represents the subcarrier number of the ℓ-th link channel sample. The sample order for each subcarrier is then shuffled to construct a matrix $\boldsymbol{H}^{\mathrm{sf}}$, which is subsequently concatenated with the original matrix $\check{\boldsymbol{H}}^{\mathrm{S,L}}$ along the sample dimension in a randomized manner to form $\boldsymbol{H}^{\mathrm{ct}}$. Finally, the method from Section 3.1.1 is applied to process $\boldsymbol{H}^{\mathrm{ct}}$. After restoring the slow-time sequence of the samples corresponding to original matrix, the resulting $\tilde{\boldsymbol{H}}^{\mathrm{S,L}}$ is the CSI with the interference completely removed. The details are provided in the **Algorithm 1**. To verify it, we take 50 CSI samples and compare the residual offsets before and after disentanglement in Figure 6(d); the results demonstrate that the residual offsets have been effectively eliminated.

3.2 Bandwidth Reconstruction

Given the well calibrated per-link CSIs (with interference disentangled and removed), these CSIs are ready to help expanding sensing bandwidth. To this end, we propose an optimization model that extracts frequency-independent components from multiple channel (CSI) samples and extends them to unseen channels. This method exploits the fact that multi-link channel samples enable an optimization objective through mutual validation, facilitating the extraction of frequency-independent information as the foundation for bandwidth expansion via samples from only a few channels.

Assuming the frequency-independent component $\Xi = F(\tilde{\boldsymbol{H}}_\ell^{\mathrm{S,L}}, f_\ell^{\mathrm{c}})$, where f_ℓ^{c} denotes the carrier frequency of the ℓ-th channel link, and F is a nonlinear mapping designed based on a modulation module [44] to remove the frequency-dependent information from the CSI $\tilde{\boldsymbol{H}}_\ell^{\mathrm{S,L}}$:

$$F\left(\tilde{\boldsymbol{H}}_\ell^{\mathrm{S,L}}, f_\ell^{\mathrm{c}}\right) = \mathrm{ReLU}\left(w * \tilde{\boldsymbol{H}}_\ell^{\mathrm{S,L}} * \gamma + \beta\right), \tag{7}$$

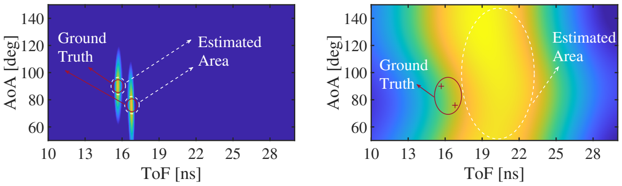
where $*$ is convolution operator, $\gamma = w_\gamma \mathrm{e}^{-\mathrm{i}2\pi \frac{f_\ell^{\mathrm{c}}}{\mathrm{c}}} + b_\gamma$, and $\beta = w_\beta \mathrm{e}^{-\mathrm{i}2\pi \frac{f_\ell^{\mathrm{c}}}{\mathrm{c}}} + b_\beta$. Here we leverage a simple (single layer) neural network to model F, where the terms w and b are the weight matrix and bias to be optimized, respectively. Ξ can then be generalized to another channel with the carrier frequency f_J^{c} as: $\tilde{\boldsymbol{H}}_J^{\mathrm{S,L}} = G(\Xi, f_J^{\mathrm{c}})$, where G shares the same structure as F but operates with different parameters. Since multiple channel samples with different carrier frequencies mutually validate the quality of the generated samples, the optimization problem is formulated as:

$$\min_{F,G} \sum_{\ell=1}^{L} \sum_{J=1}^{L} \left\| G\left(F\left(\tilde{\boldsymbol{H}}_\ell^{\mathrm{S,L}}, f_\ell^{\mathrm{c}}\right), f_J^{\mathrm{c}}\right) - \tilde{\boldsymbol{H}}_J^{\mathrm{S,L}} \right\|_2^2, \tag{8}$$

where L represents the number of channel links. With unequal bandwidths across links, the larger-bandwidth link can be logically divided into multiple virtual links to match the narrower one. Using gradient descent [48], F and G are solved to enable the extraction of feature components and the generation of unseen channel samples. As this optimization objective involves taking CSI difference across two links, inputting misaligned CSI samples (i.e., without the calibration in Section 3.1) would severely deviate the solution.

Although the above model can theoretically generate arbitrary unseen channel samples, the extracted Ξ components of different links inevitably contain errors in practice. Therefore, we generate CSIs only between the available channel samples and fine-tune them using additional generated overlapping CSIs. Specifically, for each pair of available channel samples, the Ξ components extracted from each are used to generate CSIs between them, extending toward each other until the bandwidth of the overlapping region exceeds 1/3 of the gap bandwidth. The phase and magnitude of both groups of generated CSIs are then adjusted based on the overlapping regions [66, 69] to ensure a smooth connection, yielding the bandwidth reconstruction result. Additionally, since reconstructing GHz-level bandwidth would involve thousands of subcarriers [1], spline interpolation is applied to combine multiple subcarriers, reducing the data size and alleviating computation burdens on downstream sensing applications.

In Figure 7(a), we present the localization results achieved by μCeiver-Fi with two channels at 2452 MHz and 5190 MHz and a bandwidth of 40 MHz for each. Compared with Figure 6(c), the results clearly show that the two subjects are successfully distinguished, confirming the high resolution and further validating the excellent alignment of channel samples. Furthermore, Figure 7(b) shows the results obtained by HiSAC's compressed sensing algorithm [43] upon our cleansed channel samples; the evident ambiguity in the estimated locations highlights the superiority of μCeiver-Fi in bandwith expansion on commodity multi-link Wi-Fi.



(a) Our algorithm. (b) HiSAC.

Figure 7: Comparing differently reconstructed CSIs.

3.3 Fine-Granularity Sensing Adapter

In this section, we propose a spectral representation for downstream tasks; it is extracted through a neural network that employs a model-based learning strategy exploiting the spectral representation. We then demonstrate its adaptability using the challenging task of Wi-Fi human pose estimation.

3.3.1 Sensing Spectrum Design. To enable ubiquitous sensing, it is imperative to convert CSI samples containing fine-granularity subject information into a general representation. As explicit functions can hardly realize this conversation, we resort to a neural network for this purpose. Thanks to the model-based learning framework proposed in [30], we may leverage a similar network design (or other designs for regression tasks) and the corresponding training strategy; we thus focus to the design of a spectral representation as both network output and training labels.

We consider a 3D continuous spectrum represented as three 2D spectra (similar to the "channel" concept in deep learning [60]), where each 2D spectrum is defined by two "location" parameters (x, y) and an amplitude $\aleph^{x,y}$ to represent a certain state (e.g., likelihood, strength, speed, or even shape). As this data structure acts as the network output, it should handle the bias-variance dilemma [40] inherent to statistical inference. Consequently, state information concerning the p-th subject should be characterized via a *region* centered around $(x_p, y_p, \aleph_p^{x,y})$, rather than concentrating only to $(x_p, y_p, \aleph_p^{x,y})$. Moreover, as labels for model-based training, a function should be in place to derive the amplitudes of the p-th region from known physical models or data. This function is typically defined as a variational transformation that adjusts the mathematical fuzziness based on the Euclidean distance from the p-th subject. Taking the ToF-AoA spectrum (in Figures 3, 6, and 7) as an example, the person (subject) location is indicated by $\{x, y\}$ as it is only a 2D spectral representation, while $\aleph_p$ can be realized by MUSIC that transforms CSIs to ToF-AoA spectrum [30].

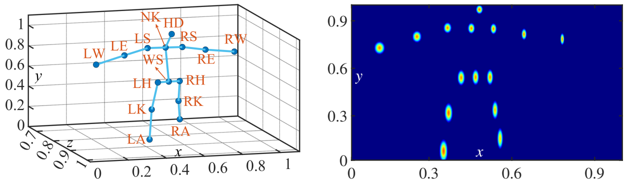
3.3.2 Adapter for Wi-Fi Human Pose Estimation. Since applying the spectrum to various human sensing tasks has been fully illustrated in [30] (also in earlier figures), we hereby focus on a more challenging task, namely Wi-Fi 3D human pose estimation. This task serves as an ideal benchmark for assessing Wi-Fi sensing performance, as it requires both high accuracy and resolution to distinguish multiple keypoints on a subject. Since the sensing subjects in Wi-Fi pose estimation are *keypoints* on, for example, a person [52, 59], the physical coordinates (x_p, y_p, z_p) of these keypoints are obtained via, e.g., OpenPose [4], from two correlated photos. Then the state information around (x_p, y_p) is given by $\aleph_p^{x,y}$ as

$$\aleph_p^{x,y} = z_p \cdot \hat{\aleph}_p^{x,y} + \eta, \tag{9}$$

where η represents Gaussian random noise; it allows the neural network to convert CSIs with unseen variations to different spectra [30], and $\hat{\aleph}_p^{x,y}$ is defined as

$$\hat{\aleph}_p^{x,y} = \exp\left(\frac{-\left(x - x_p\right)^2}{2\left(\sigma_p^x(1 - |x - x_p|)\right)^2} + \frac{-\left(y - y_p\right)^2}{2\left(\sigma_p^y(1 - |y - y_p|)\right)^2} \right),$$

with $\sigma_p^x = \sigma \cdot (|x_p - \max(x)|)$, $\sigma_p^y = \sigma \cdot (|y_p - \max(y)|)$, and σ being the standard deviation. Similarly, we may derive another two spectra to obtain the "three-channel" 3D spectrum $\aleph_p = \{\aleph_p^{x,y}, \aleph_p^{y,z}, \aleph_p^{x,z}\}$. Essentially, Eqn. (9) represents customized "perspective" projections from the 3D formation of a subject onto three planes in the spectrum. We adopt the widely validated keypoint definitions [4], describing the human body by 15 keypoints: head (HD), neck (NK), left and right shoulders (LS and RS), left and right elbows (LE and RE), left and right wrists (LW and RW), waist (WS), left and right hips (LH and RH), left and right knees (LK and RK), and left and right ankles (LA and RA), as shown in Figure 8(a). Then Figure 8(b) shows an example of the spectrum on the (x, y)-plane that integrates individual subject spectra derived via Eqn. (9) under certain states. This spectrum makes μCeiver-Fi superior to existing Wi-Fi pose estimation in that, besides locating the keypoints, certain states (e.g., respiration) can be further extracted from the spectrum [30].

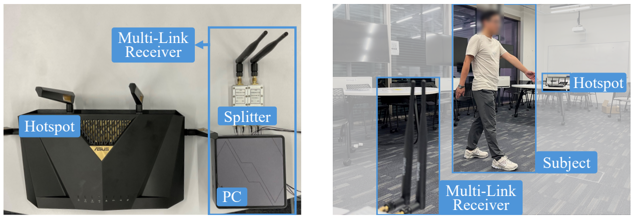


(a) Human keypoint coordinates. (b) Spectrum.

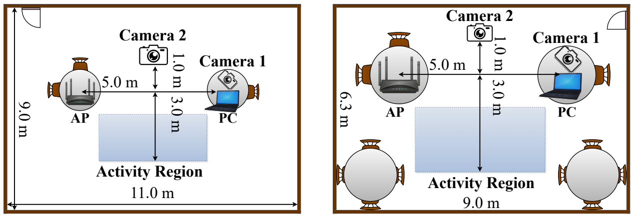
Figure 8: Labels for the Wi-Fi pose estimation network.

4 PROTOTYPE AND EXPERIMENT SETUP

We build μCeiver-Fi prototype using an ASUS RT-BE88U [3] hotspot and a mini PC equipped with an R7-7840H AMD CPU and 16 GB RAM, as shown in Figure 9(a). While ASUS RT-BE88U is capable of multi-link operation, we have no access to CSI on IEEE 802.11be NICs [1]. We thus use two Intel AX210 NICs [11] on the mini PC to emulate a multi-link receiver: sharing two antennas via power splitters, they become equivalent to a multi-link receiver built with a single NIC. The hotspot chip currently supports simultaneous connections only to the 2.4 GHz and 5 GHz bands; it is equipped with four antennas, with two assigned to each band. The mini PC streams 4K video from the hotspot to get a traffic rate of approximately 200 packets/s, and it captures CSI samples via the PicoScenes [23]. Two cameras are employed for human image acquisition: one positioned on the mini PC at a 45° angle to the LoS path, and the other close to the midpoint of the LoS path but perpendicular to it. The raw CSI is then processed in MATLAB, following the procedure outlined in Sections 3.1-3.2 to obtain wideband signals. Simultaneously, the ground truth images are converted into spectrum labels, as detailed in Section 3.3, and aligned with the CSI data based on timestamps. Since our contribution focuses on CSI processing and label design, the validated neural network from UWB-Fi [30] is implemented in PyTorch to transform wideband CSIs into spectra. A thresholding technique is then applied to extract peaks from the spectra, which are traced from two near-ground points, namely LA and RA, to infer their corresponding keypoints based on human body structure, thereby enabling pose estimation.



(a) μCeiver-Fi prototype. (b) A testing case.



(c) Rehearsal room layout. (d) Meeting room layout.

Figure 9: Photographs of (a) μCeiver-Fi prototype and (b) a testing scenario, and layouts of (c) rehearsal room and (d) meeting room.

We collect over 40 hours of CSI data from 12 (human) subjects performing daily activities in two different scenarios to evaluate our μCeiver-Fi system. The subjects, consisting of eight males and four females aged 20 to 30, have heights ranging from 160 cm to 185 cm. The two scenarios are a rehearsal room (RR) and a meeting room (MR). Each subject performs a series of postures, including standing (ST), hands up (HU), hands open (HO), pointing (PT), walking (WK), and sitting down (SD), as shown in Figure 9(b). Subjects are instructed to move within a range of 1.5 m to 3 m from the LoS path, as shown in Figures 9(c) and 9(d). Besides, the LoS path length is varied from 3 m to 7 m to cover a broad range of scenarios. We establish simultaneous connections on channels at 2452MHz and 5210MHz, using bandwidths of 40MHz and 80MHz, respectively. The dataset is segmented so that CSIs within a 30ms period are deemed as one *sample* for pose estimation, with multiple consecutive samples for other sensing tasks (e.g., respiration monitoring). All experiments strictly adhere to the IRB guidelines of our institute.

The data from one subject in the rehearsal room performing activities near a 2 m distance from the LoS path, with a LoS path length of 5m, is used as the training set for the neural network, while the remaining data are used for testing. Furthermore, we implement experiments based on acknowledged proposals [22, 47, 63, 64] for comparison. WiPose [22] extracts the body-coordinate velocity profile (BVP) from CSI and maps it to quaternions using a neural network for 3D pose estimation, designated as **Baseline 1**. A series of approaches proposed in [47, 63, 64] primarily transform CSI into 2D AoA spectra and employ neural networks to estimate keypoints, referred to as **Baseline 2**. For a fair comparison, both baselines are evaluated using data collected with Intel AX210 NICs equipped with two antennas and operating on an 80 MHz bandwidth. We refrain from comparing with HiSAC [43] and SiWis [52] due to their distinctive hardware.

To comprehensively evaluate pose estimation performance, average precision (AP) [33] is introduced as a similarity metric alongside localization error. The AP metric is an extension of the object keypoint similarity (OKS):

$$OKS = \sum_p \exp(-\mathcal{D}_p^2/2s^2\Upsilon_p^2)\mathbb{I}(v_p > 0)(\sum_p \mathbb{I}(v_p > 0))^{-1}, \quad (10)$$

where $\mathcal{D}_p$ represents the distance between the p-th predicted keypoint and the ground truth. The scale factor s is the detection box diagonal length, Υ is a normalization factor for different keypoints, typically treated as a constant [33], and $\mathbb{I}$ denotes the indicator function, which equals 1 if the visibility parameter $v > 0$ and 0 otherwise. Then, AP is defined as:

$$AP = P^{-1}\sum_p \mathbb{I}(OKS_p > a), \quad (11)$$

where P denotes the total number of keypoints. The mean AP (mAP) is obtained by averaging the AP values corresponding to different thresholds of a, where $a = [0.5 : 0.05 : 0.95]$.

5 EVALUATIONS

In this section, we conduct a comprehensive evaluation of μCeiver-Fi through Wi-Fi human pose estimation, a widely studied yet challenging application. We begin by comparing μCeiver-Fi with two baseline methods in visualization quality and overall performance, highlighting its advantages in Wi-Fi sensing. The generalization capability of μCeiver-Fi is then assessed by performing respiration sensing over pose estimation and by analyzing the impact of subject variations, environments, and experiment parameters on results. Finally, ablation studies are performed to validate the necessity and effectiveness of the proposed algorithmic modules.

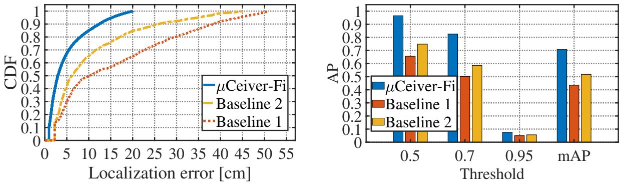
5.1 Visualized Comparisons

We select frames of six representative actions, as shown in Figure 10(a), and leverage OpenPose [4] to convert them into 3D keypoint coordinates as ground truths, which are visualized in Figure 10(b). The visualization results of μCeiver-Fi and two baselines are presented in Figure 10(c) to (e), with different colors distinguishing keypoints complying with corresponding ground truths from those not. It is clear that the results by μCeiver-Fi accurately depict the subjects' postures: even with a distance of about 15 cm, the keypoints HD and NK can be clearly distinguished, indicating the GHz-level effective sensing bandwidth. Moreover, all keypoint coordinates exhibit very small error compared with the ground truths, with only slight deviations observed along the z-axis of the knee joint under the ST, HU, HO, and PT postures, while barely perceptible under the WK and SD postures.

The pose estimation results of Baseline 1 shown in Figure 10(d) indicate that, while they roughly capture the subjects' actions, the keypoints exhibit significant deviations from the ground truth, leading to noticeable distortions. Beyond the commonly observed leg coordinate errors, Baseline 1 also introduces deformations in the head and arms during the ST and HU actions, highlighting its difficulty in accurately identifying closely spaced keypoints. The pose estimation results of Baseline 2, displayed in Figure 10(e), reveal that distortions primarily occur in the legs and head. Although its errors are lower than those in Baseline 1, the overall pose estimation quality remains inferior to our μCeiver-Fi. Moreover, Baseline 2 introduces significant errors in the arm region during the WK posture, suggesting it struggles with involuntary body jitter, which may degrade the quality of both CSIs and training labels. Overall, μCeiver-Fi outperforms both baselines, demonstrating its advantage of the wide effective sensing bandwidth in handling closely spaced keypoints and adverse scenarios.

5.2 Overall Performance

We now quantitatively evaluate the overall performance of μCeiver-Fi against baselines. The average localization error for all keypoints of a subject is first assessed, with the cumulative distribution functions (CDFs) of μCeiver-Fi and the two baselines shown in Figure 11(a). μCeiver-Fi has a median error around 3 cm, the 90th percentile error around 12 cm, and the maximum error below 20 cm. In contrast, Baseline 1 yields respective values of 10cm, 38cm, and 50cm, while those with Baseline 2 are 7cm, 26cm, and 45cm, respectively. These results confirm that μCeiver-Fi consistently achieves higher keypoints localization accuracy than the other two baselines. To further assess overall pose estimation performance, we present the AP values for threshold values $a = [0.5, 0.7, 0.95]$ and the mAP under $a = [0.5 : 0.05 : 0.95]$ in Figure 11(b). Although all three methods exhibit performance degradation as the threshold increases, μCeiver-Fi maintains an AP > 0.8 even at $a = 0.7$ and achieves the highest mAP of 0.71, clearly demonstrating its minimal distortions.



(a) Localization error.

(b) AP value.

Figure 11: Overall performance of pose estimation.

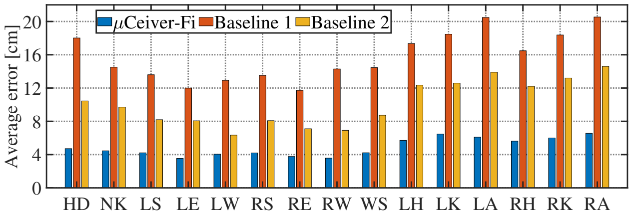


Figure 12: Average localization errors of each keypoint recognized by μCeiver-Fi, Baseline 1, and Baseline 2.

To gain deeper insight into the pose estimation details, we analyze the average localization error for each keypoint, as shown in Figure 12. The results reveal that μCeiver-Fi consistently achieves lower errors across all keypoints. A closer examination reveals that the statistical trends in Figure 12 generally align with the visualization results in Figure 10: both baselines exhibit lower localization accuracy for head, neck, and legs, while the torso region shows relatively higher accuracy. This is mainly due to the small size and close proximity of these keypoints, as the baselines lack sufficient range resolution to ensure their accurate identification. Despite slight fluctuations in leg localization accuracy, μCeiver-Fi achieves satisfactory 3D human pose estimation results, highlighting the benefits of its GHz-level effective sensing bandwidth.

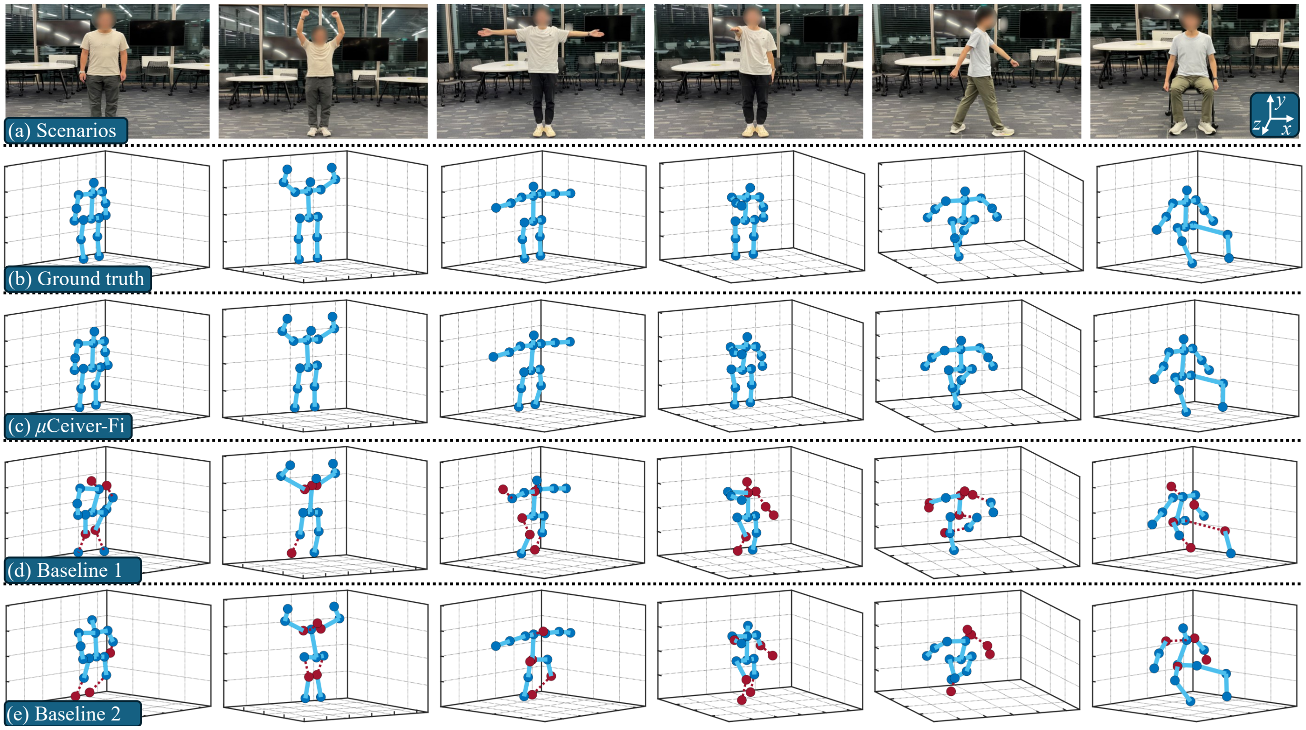
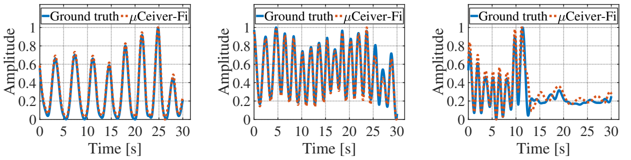


Figure 10: Human postures (from left to right) ST, HU, HO, PT, WK, SD. (a) Testing scenarios, (b) ground truth, and Wi-Fi human pose estimation results of (c) μCeiver-Fi, (d) Baseline 1, and (e) Baseline 2.



(a) Slow respiration. (b) Fast respiration. (c) Irregular respiration.

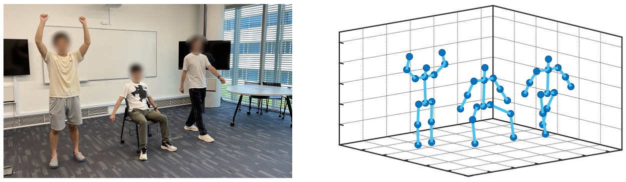
Figure 13: Respiration waveforms recovered from (a) slow, (b) fast, and (c) irregular patterns.

5.3 Generalization Evaluation

We hereby showcase the respiration sensing and analyze the impact factors of μCeiver-Fi using the mAP metric.

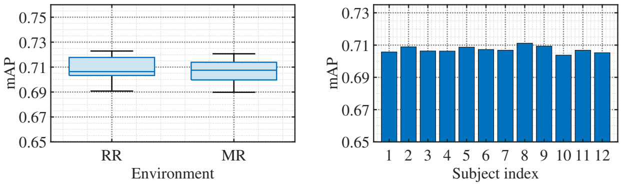
5.3.1 Respiration Sensing. Although μCeiver-Fi significantly surpasses the baselines in Wi-Fi pose estimation, this task alone may not fully reflect the advantage of its GHz-level sensing bandwidth. Therefore, we further validate a unique capability of μCeiver-Fi: achieving fine-granularity sensing tasks by leveraging the spectra. When a subject takes the SD posture, μCeiver-Fi is able to further extract its respiration waveforms, using several consecutive imaging frames and focusing on the region close to the NK and WS keypoints. We leverage MoRe-Fi [74] to recover the respiration waveform from these frames and acquire ground truth using a NeuLog chest strap [41]. The results show that the average respiration rate error is approximately 0.1 bpm. Figure 13 further visualizes that the waveforms under slow, fast, and irregular breathing patterns are accurately recovered, with only minor deviations occurring when the subject holds breath, i.e. after 13 s in Figure 13(c), due to the stillness of the chest and abdomen. In fact, μCeiver-Fi can sense hand gestures and gaits via the same approach, i.e., using a sequence of imaging frames but focusing on different regions.

5.3.2 Multiple Subjects. To evaluate the performance of μCeiver-Fi in multi-person pose estimation tasks, three subjects are instructed to perform different postures, as shown in Figure 14(a). The visualized results in Figure 14(b) demonstrate that the three subjects are clearly distinguished, and all their keypoints are accurately localized, leading to satisfactory pose estimation. This observation indicates that, due to the sufficient range resolution and diversity information provided by the large effective bandwidth, our μCeiver-Fi is capable of handling complex scenarios.



(a) Scenario. (b) Pose estimation result.

Figure 14: Multi-person pose estimation.

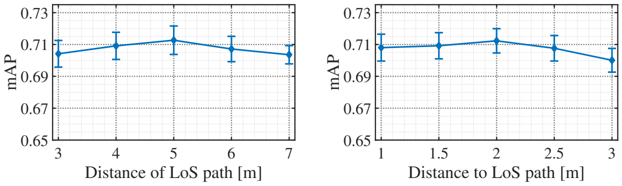


(a) Impact of environment. (b) Impact of subject.

Figure 15: Impact of (a) environment and (b) subject.

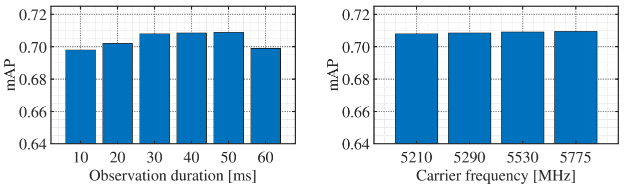
5.3.3 Environment and Subject. To assess the impact of the environment on pose estimation, the mAP values of subjects in RR and MR scenarios are examined. According to Figure 15(a), the environment has only minor impact on μCeiver-Fi's pose estimation performance, with only a slight decline observed in the cluttered MR setting filled with tables. This robustness primarily stems from the high resolution of μCeiver-Fi, which effectively separates the subject from the background. The mAP values for individual subjects are then examined in Figure 15(b), revealing only minor fluctuations. Subject 8 achieves the highest accuracy, while Subject 10 performs slightly worse, with further analysis attributing the discrepancy to body size differences: Subject 10 is about 20 cm shorter than Subject 8. Despite this considerable variation in physique, no significant performance degradation is observed, thus demonstrating the strong generalizability of μCeiver-Fi across environment and subject.

5.3.4 Distance. We also examine the impact of LoS path distance by varying it from 3 m to 7 m. The results shown in Figure 16(a) indicate that LoS path distance has a minor effect on pose estimation accuracy, with a slight improvement observed at 5 m compared to other cases. Further analysis suggests that excessively short distances may cause the strong LoS path signal to overshadow subject-reflected components, while overly long distances weaken the reflected signal power, ultimately degrading performance. We further study the impact of the subject's distance to the LoS path on Wi-Fi human pose estimation. The results in Figure 16(b) exhibit a similar trend: being too close yields higher interference from LoS (thus lower mAP), yet moving further reduces the power of the reflected signals (again lowing mAP). Nevertheless, the mAP values remain around 0.7, demonstrating satisfactory performance despite the variations in distance.



(a) LoS path distance. (b) Subject's distance to LoS path.

Figure 16: Impact of distance.

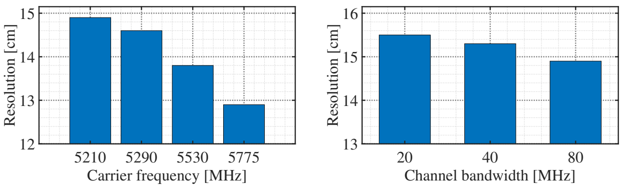


(a) Impact of observation duration. (b) Impact of channel spacing.

Figure 17: Impact of duration and spacing.

5.3.5 Duration and Channel. To analyze the impact of observation duration on pose estimation, we obtain mAP under durations ranging from 10 ms to 60 ms. The results in Figure 17(a) indicate that, while increasing the observation duration enhances pose estimation performance up to around 30 ms, further extending it leads to a decline in accuracy. Further analysis reveals that the degradation is primarily due to the excessive duration exceeding the channel coherence time, thus reducing signal stability. The results highlight the advantage of μCeiver-Fi over channel hopping in capturing sufficient channel samples within the shortest possible time, as discussed in Section 2.2. To assess the impact of the spacing between two channels on pose estimation, the channel index in the 5 GHz band is gradually adjusted to test 5210 MHz, 5290 MHz, 5530 MHz, and 5775 MHz channels, with the channel bandwidth fixed at 80 MHz. As shown in Figure 17(b), the pose estimation performance remains stable as the channel spacing varies, since the pose estimation task has already acquired sufficient bandwidth, and further increases do not provide significant advantages. To reduce computation burden, the channel with a carrier frequency of 5210 MHz is selected in our study, with the option to adjust it for other tasks requiring higher resolution.

5.3.6 Range Resolution. Since pose estimation cannot precisely quantify range resolution, we employ two pendulums to analyze the impact of channel spacing and bandwidth. As shown in Figure 18(a), with the 5 GHz channel bandwidth fixed at 80 MHz, increasing the spacing improves the range resolution from approximately 15 cm to 13 cm. This result demonstrates the strong potential of our μCeiver-Fi to reconstruct a larger effective sensing bandwidth once the 6 GHz band becomes available for multi-link receivers. In addition, we fix the carrier frequency of the 5 GHz channel around 5210 MHz and vary its bandwidth. As illustrated in Figure 18(b), the range resolution improves with increasing channel bandwidth due to the capture of more information.



(a) Impact of channel spacing. (b) Impact of bandwidth.

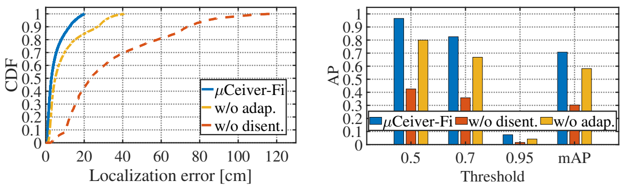
Figure 18: Impact on range resolution.

5.4 Ablation Study

To evaluate the importance of each algorithm module in our μCeiver-Fi, pose estimation performance is analyzed after sequentially removing them, as shown in Figure 19.

First, instead of applying the disentanglement method proposed in Section 3.1, interference is calibrated solely through HiSAC's search algorithm [43]. Meanwhile, given that our bandwidth reconstruction algorithm is designed to work with the disentanglement method, this module is also replaced with that of HiSAC to obtain wideband CSIs, with the corresponding pose estimation accuracy shown in red. The results indicate a significant performance degradation, with the median and 90th percentile localization errors around 23 cm and 74 cm, while the maximum error exceeds 1 m and the mAP value drops to around 0.3, which is far from usable considering the typical body size of an adult. Further analysis reveals that the presence of residual offsets leads to severe distortions, while their randomness imposes significant fitting pressure on the pose estimation neural network. Besides, the bandwidth reconstruction also fails with two-channel samples as HiSAC has access to four or five channels. These results fully demonstrate the effectiveness of μCeiver-Fi upon commodity multi-link Wi-Fi devices.

The sensing adapter described in Section 3.3 is then removed, and the spectral parameters used as training labels are replaced with common 3D coordinate tuples. The corresponding accuracy is represented in yellow, and the results indicate that, although slightly surpassing the baselines due to the wide bandwidth of input CSIs, a general decline in performance exists compared with μCeiver-Fi. This is attributed to the limited fitting capability of the neural network model: directly mapping the input signals to tuples introduces inevitable errors due to the bias-variance dilemma [40], while our method mitigates this by offloading bias to regions unrelated to the subject's keypoints, thus preserving pose estimation accuracy in areas of interest. More importantly, the spectra can then be leveraged to embed state information for serving other fine-granularity sensing tasks.



(a) Localization error. (b) AP value.

Figure 19: Impact of algorithms in μCeiver-Fi.

6 RELATED WORK AND DISCUSSION

In this section, we review recent proposals related to our work and discuss the future development of μCeiver-Fi.

CSI Calibration. Phase offsets and noise pose significant challenges in Wi-Fi sensing and substantially affect its performance, thereby necessitating calibration. ArrayTrack [68] estimates offsets by leveraging specially designed cables, but it cannot adapt to offset changes after a NIC restart. To enhance practicality, SpotFi [25] uses linear regression to eliminate offsets. Chronos [57] and Widar2.0 [46] achieve this by using conjugates of CSI between the uplink and downlink, and between different antennas, respectively. In addition, inspired by the monostatic configuration of FMCW radar (which eliminates the need for complex offset calibration), ISACoT [7] implements a monostatic Wi-Fi system on the SDR platform for sensing. Another common approach [13, 43] is to employ search algorithms to determine the compensation offsets that produce an accurate LoS path estimation. However, this method leaves residual offsets and cannot align multiple channel samples.

Sensing Bandwidth Expansion. Wi-Fi bandwidth expansion is crucial for achieving the high resolution needed for fine-granularity sensing. ToneTrack [69] and Splicer [66] collect a large number of overlapping channel samples to calibrate adjacent signals for bandwidth reconstruction, while Chronos [57] gathers continuous channel samples and process the 0-th subcarrier for device localization. To improve the sampling efficiency and address the issue that not all channels are available, UWB-Fi [30] collects CSIs leveraging a fast channel hopping scheme and converts discrete channel samples into a ToF-AoA spectrum for ubiquitous sensing. Frequency hopping is also adopted by RFind [36] and RFChord [32] for RFIDs that require no calibration, while HiSAC [43] collects concurrent channel samples using multi-link mmWave SDRs. However, all these systems cannot be readily deployed on commodity devices.

Wi-Fi Human Pose Estimation. Wi-Fi human pose estimation effectively addresses privacy and low-light concerns, emerging as a promising field. Wision [19] leverages 2D-FFT to separate multipath reflections for 2D pose estimation through azimuthal and elevation angles, while WiSIA [26] combines 2D-IFFT with cGAN to achieve the same purpose. Wifract [42] further enhances accuracy by employing carefully designed antenna arrays and diffraction theory to identify subject edges. In 3D pose estimation, WiPose [22] transforms CSIs into 3D keypoint data via an end-to-end neural network, while GoPose [47] and Wi-Mesh [63] utilize multiple devices to estimate the 2D AoA for human body construction. Additionally, SiWiS [52] employs multiple directional antennas and an FPGA to build a monostatic [7] Wi-Fi sensing system for pose estimation.

Discussions. Since the latest NICs compliant with the IEEE 802.11be standard [20] do not yet support publicly available CSI access, we implement μCeiver-Fi prototype using two Intel AX210 NICs [11], providing a practical and effective solution under current hardware constraints. These NICs utilize power splitters to share common antennas, functionally equating to the system where one NIC integrates two links at the MAC layer. In the future, driven by the trend toward open CSI access, we plan to conduct a comprehensive evaluation on a system built with one NIC to investigate potential unknown factors, which are expected to enhance μCeiver-Fi performance through better inter-module synchronization. Moreover, although two NICs are used, μCeiver-Fi fundamentally differs from existing multi-device systems [18, 54] by simultaneously connecting two links operating at different frequencies, thereby enabling bandwidth expansion. Meanwhile, μCeiver-Fi can be further extended by deploying several multi-link receivers, offering significant promise for further performance enhancement. Beyond the demonstrated tasks, μCeiver-Fi shows strong prospects for a range of applications, including gait analysis, gesture recognition, and fall detection. It is also expected to be further evaluated in more adverse scenarios, such as those involving occlusions and rich multipath environments. Last but not least, co-existing with other co-channel communication systems [27, 31, 70] should be handled in the future.

7 CONCLUSION

We have introduced μCeiver-Fi, a system leveraging a commodity multi-link receiver to explore spectrum sources for effective sensing bandwidth expansion. μCeiver-Fi first employs a disentanglement algorithm to fully eliminate interference and align channel samples from different links. A bandwidth reconstruction model is then built to transform multiple channel samples into a wideband signal. Finally, a general spectral representation is developed for μCeiver-Fi as an adapter to gain versatility across various applications. Based on these, we have implemented μCeiver-Fi prototype and evaluated it in a Wi-Fi human pose estimation task. The results evidently confirm μCeiver-Fi's capability in fine-granularity sensing. The codes accompanying this paper are available via https://github.com/DeepWiSe888/muCeiver-Fi.

ACKNOWLEDGEMENT

This research is supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-GC-2023-006), the National Research Foundation Singapore and Infocomm Media Development Authority under its Future Communications Research & Development Programme, and MOE Tier 1 grant RG16/22.

REFERENCES

[1] 2024. IEEE Approved Draft Standard for Information Technology–Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks–Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment: Enhancements for Extremely High Throughput (EHT). *IEEE P802.11be/D7.0, August 2024* (2024), 1–1089.

[2] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D Tracking via Body Radio Reflections. In *Proc. of the 11th USENIX NSDI.* 317–329.

[3] ASUSTeK Computer Inc. 2024. User Guide: RT-BE88U BE7200 Dual Band Wi-Fi Router. https://dlcdnets.asus.com/pub/ASUS/wireless/RT-BE88U/E23805_RT-BE88U_UM_WEB.pdf?model=RT-BE88U/.

[4] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-person 2D Pose Estimation Using Part Affinity Fields. In *Proc. of the 30th IEEE/CVF CVPR.* 7291–7299.

[5] Yan Chen, Xiang Su, Yang Hu, and Bing Zeng. 2019. Residual Carrier Frequency Offset Estimation and Compensation for Commodity WiFi. *IEEE Transactions on Mobile Computing* 19, 12 (2019), 2891–2902.

[6] Zhe Chen, Zhongmin Li, Zhang Xu, Guorong Zhu Zhu, Yuedong Xu, Jie Xiong, and Xin Wang. 2017. AWL: Turning Spatial Aliasing From Foe to Friend for Accurate WiFi Localization. In *Proc. of the 13th ACM CoNEXT.* 238–250.

[7] Zhe Chen, Tianyue Zheng, Chao Hu, Hangcheng Cao, Yanbing Yang, Hongbo Jiang, and Jun Luo. 2023. ISACoT: Integrating Sensing with Data Traffic for Ubiquitous IoT Devices. *IEEE Communications Magazine* 61, 5 (2023), 98–104.

[8] Zhe Chen, Tianyue Zheng, and Jun Luo. 2021. Octopus: A Practical and Versatile Wideband MIMO Sensing Platform. In *Proc. of the 27th ACM MobiCom.* 601–614.

[9] Zhe Chen, Guorong Zhu, Sulei Wang, Yuedong Xu, Jie Xiong, Jin Zhao, Jun Luo, and Xin Wang. 2021. M^3: Multipath Assisted Wi-Fi Localization with a Single Access Point. *IEEE Trans. on Mobile Computing* 20, 2 (2021), 588–602.

[10] Krishna Chintalapudi, Bozidar Radunovic, Vlad Balan, Michael Buettener, Srinivas Yerramalli, Vishnu Navda, and Ramachandran Ramjee. 2012. WiFi-NC: WiFi Over Narrow Channels. In *Proc. of the 9th USENIX NSDI.* 43–56.

[11] Intel Corporation. 2023. Intel® Wi-Fi 6 AX201. https://www.intel.sg/content/www/xa/en/products/sku/130293/intel-wifi-6-ax201-gig/specifications.html. Online; accessed 25 March 2023.

[12] Shuya Ding, Zhe Chen, Tianyue Zheng, and Jun Luo. 2020. RF-Net: A Unified Meta-Learning Framework for RF-enabled One-Shot Human Activity Recognition. In *Proc. of the 18th ACM SenSys.* 517–530.

[13] Jon Gjengset, Jie Xiong, Graeme McPhillips, and Kyle Jamieson. 2014. Phaser: Enabling Phased Array Signal Processing on Commodity WiFi Access Points. In *Proc. of the 20th ACM MobiCom.* 153–164.

[14] Gene H Golub and Charles F Van Loan. 2013. *Matrix Computations.*

[15] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool Release: Gathering 802.11n Traces with Channel State Information. *ACM SIGCOMM Comput. Commun. Rev.* 41, 1 (2011), 53.

[16] Yinghui He, Jianwei Liu, Mo Li, Guanding Yu, Jinsong Han, and Kui Ren. 2023. SenCom: Integrated Sensing and Communication with Practical WiFi. In *Proc. of the 29th ACM MobiCom.* 1–16.

[17] Peter Hillyard, Anh Luong, Alemayehu Solomon Abrar, Neal Patwari, Krishna Sundar, Robert Farney, Jason Burch, Christina Porucznik, and Sarah Hatch Pollard. 2018. Experience: Cross-technology Radio Respiratory Monitoring Performance Study. In *Proc. of the 24th ACM MobiCom.* 487–496.

[18] Jingzhi Hu, Tianyue Zheng, Zhe Chen, Hongbo Wang, and Jun Luo. 2023. MUSE-Fi: Contactless MUti-person SEnsing Exploiting Near-field Wi-Fi Channel Variation. In *Proc. of the 29th ACM MobiCom.* 75:1–15.

[19] Donny Huang, Rajalakshmi Nandakumar, and Shyamnath Gollakota. 2014. Feasibility and Limits of Wi-Fi Imaging. In *Proc. of the 12th ACM SenSys.* 266–279.

[20] Intel Corporation. 2024. Intel Wi-Fi 7 BE201 (Fillmore Peak 2) Product Brief. https://www.intel.com/content/www/us/en/content-details/823795/intel-wi-fi-7-be201-fillmore-peak-2-product-brief.html. Online; accessed 27 May 2024.

[21] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, Wenyao Xu, and Lu Su. 2018. Towards Environment Independent Device Free Human Activity Recognition. In *Proc. of the 24th ACM MobiCom.* 289–304.

[22] Wenjun Jiang, Hongfei Xue, Chenglin Miao, Shiyang Wang, Sen Lin, Chong Tian, Srinivasan Murali, Haochen Hu, Zhi Sun, and Lu Su. 2020. Towards 3D Human Pose Construction Using WiFi. In *Proc. of the 26th MobiCom.* 1–14.

[23] Zhiping Jiang, Tom H. Luan, Xincheng Ren, Dongtao Lv, Han Hao, Jing Wang, Kun Zhao, Wei Xi, Yueshen Xu, and Rui Li. 2021. Eliminating the Barriers: Demystifying Wi-Fi Baseband Design and Introducing the PicoScenes Wi-Fi Sensing Platform. *IEEE Internet of Things Journal* (2021), 1–21.

[24] Chitra R. Karanam, Belal Korany, and Yasamin Mostofi. 2019. Tracking from One Side: Multi-person Passive Tracking with WiFi Magnitude Measurements. In *Proc. of the 18th ACM/IEEE IPSN.* 181–192.

[25] Manikanta Kotaru, Kiran Joshi, Dinesh Bharadia, and Sachin Katti. 2015. SpotFi: Decimeter Level Localization Using WiFi. In *Proc. of the 2015 ACM SIGCOMM.* 269–282.

[26] Chenning Li, Zheng Liu, Yuguang Yao, Zhichao Cao, Mi Zhang, and Yunhao Liu. 2020. Wi-Fi See It All: Generative Adversarial Network-augmented Versatile Wi-Fi Imaging. In *Proc. of the 18th ACM SenSys.* 436–448.

[27] Feng Li, Jun Luo, Gaotao Shi, and Ying He. 2013. FAVOR: Frequency Allocation for Versatile Occupancy of Spectrum in Wireless Sensor Networks. In *Proc. of the 14th ACM MobiHoc.* 39–48.

[28] Xin Li, Jingzhi Hu, and Jun Luo. 2024. Efficient Beamforming Feedback Information-Based Wi-Fi Sensing by Feature Selection. *IEEE Wireless Communications Letters* (2024).

[29] Xin Li, Jingzhi Hu, Hongbo Wang, Zhe Chen, and Jun Luo. 2025. Enabling Ultra-Wideband Wi-Fi Sensing via Sparse Channel Sampling. *IEEE Journal on Selected Areas in Communications* (2025), 1–14.

[30] Xin Li, Hongbo Wang, Zhe Chen, Zhiping Jiang, and Jun Luo. 2024. UWB-Fi: Pushing Wi-Fi towards Ultra-wideband for Fine-Granularity Sensing. In *Proc. of the 22nd ACM MobiSys.* 42–55.

[31] Li, Feng and Luo, Jun and Shi, Gaotao and He, Ying. 2017. ART: Adaptive fRequency-Temporal Co-Existing of ZigBee and WiFi. *IEEE Trans. on Mobile Computing* 16, 3 (2017), 662–674.

[32] Bo Liang, Purui Wang, Renjie Zhao, Heyu Guo, Pengyu Zhang, Junchen Guo, Shunmin Zhu, Hongqiang Harry Liu, Xinyu Zhang, and Chenren Xu. 2023. RF-Chord: Towards Deployable RFID Localization System for Logistic Networks. In *Proc. of the 20th USENIX NSDI.* 1783–1799.

[33] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. 2014. Microsoft COCO: Common Objects in Context. In *Proc. of the 13rd ECCV.* 740–755.

[34] Jian Liu, Hongbo Liu, Yingying Chen, Yan Wang, and Chen Wang. 2019. Wireless Sensing for Human Activity: A Survey. *IEEE Communications Surveys & Tutorials* 22, 3 (2019), 1629–1645.

[35] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking Vital Signs During Sleep Leveraging Off-the-Shelf WiFi. In *Proc. of the 16th ACM MobiHoc.* 267–276.

[36] Yunfei Ma, Nicholas Selby, and Fadel Adib. 2017. Minding the Billions: Ultra-Wideband Localization for Deployed RFID Tags. In *Proc. of the 23rd ACM MobiCom*. 248–260.

[37] Yongsen Ma, Gang Zhou, and Shuangquan Wang. 2019. WiFi Sensing With Channel State Information: A Survey. *ACM Computing Surveys (CSUR)* 52, 3 (2019), 1–36.

[38] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data Using t-SNE. *Journal of Machine Learning Research* 9 (2008), 2579–2605.

[39] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Abdolreza Hatamlou. 2016. Multi-verse Optimizer: A Nature-inspired Algorithm for Global Optimization. *Neural Computing and Applications* 27 (2016), 495–513.

[40] Kevin P Murphy. 2012. *Machine Learning: A Probabilistic Perspective.* MIT press.

[41] NeuLog. 2017. Respiration Monitor Belt Logger Sensor NUL-236. https://neulog.com/respiration-monitor-belt/. Online; accessed 12 February 2023.

[42] Anurag Pallaprolu, Belal Korany, and Yasamin Mostofi. 2022. Wiffract: A New Foundation for RF Imaging via Edge Tracing. In *Proc. of the 28th MobiCom*. 255–267.

[43] Jacopo Pegoraro, Jesus O Lacruz, Michele Rossi, and Joerg Widmer. 2024. HiSAC: high-resolution sensing with multiband communication signals. In *Proc. of the 22nd ACM SenSys*. 549–563.

[44] Ethan Perez, Florian Strub, Harm De Vries, Vincent Dumoulin, and Aaron Courville. 2018. FiLM: Visual Reasoning with a General Conditioning Layer. In *Proc. of the 32nd AAAI*, Vol. 32.

[45] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-Level Passive Tracking via Velocity Monitoring with Commodity Wi-Fi. In *Proc. of the 18th ACM MobiHoc*. 6:1–10.

[46] Kun Qian, Chenshu Wu, Yi Zhang, Guidong Zhang, Zheng Yang, and Yunhao Liu. 2018. Widar2.0: Passive Human Tracking with a Single Wi-Fi Link. In *Proc. of the 16th ACM MobiSys*. 350–361.

[47] Yili Ren, Zi Wang, Yichao Wang, Sheng Tan, Yingying Chen, and Jie Yang. 2022. GoPose: 3D Human Pose Estimation Using WiFi. *Proc. of the 24th ACM UbiComp* 6, 2 (2022), 1–25.

[48] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. 1986. Learning Representations by Back-Propagating Errors. *Nature* 323, 6088 (1986), 533–536.

[49] R. Schmidt. 1986. Multiple Emitter Location and Signal Parameter Estimation. *IEEE Trans. on Antennas and Propagation* 34, 3 (1986), 276–280.

[50] Xianbiao Shu, Fatih Porikli, and Narendra Ahuja. 2014. Robust Orthonormal Subspace Learning: Efficient Recovery of Corrupted Low-Rank Matrices. In *Proc. of the 27th IEEE/CVF CVPR*. 3874–3881.

[51] Elahe Soltanaghaei, Avinash Kalyanaraman, and Kamin Whitehouse. 2018. Multipath Triangulation: Decimeter-level WiFi Localization and Orientation with a Single Unaided Receiver. In *Proc. of the 16th ACM MobiSys*. 376–388.

[52] Kunzhe Song, Qijun Wang, Shichen Zhang, and Huacheng Zeng. 2024. SiWiS: Fine-grained Human Detection Using Single WiFi Device. In *Proc. of the 30th ACM MobiCom*. 1439–1454.

[53] Kun Tan, Ji Fang, Yuanyang Zhang, Shouyuan Chen, Lixin Shi, Jiansong Zhang, and Yongguang Zhang. 2010. Fine-grained Channel Access in Wireless LAN. In *Proc. of 24th ACM SIGCOMM*. 147–158.

[54] Sheng Tan, Linghan Zhang, Zi Wang, and Jie Yang. 2019. MultiTrack: Multi-user Tracking and Activity Recognition using Commodity WiFi. In *Proc. of the 37th ACM CHI*. 1–12.

[55] David Tse and Pramod Viswanath. 2005. *Fundamentals of Wireless Communication.* Cambridge university press.

[56] Jayakrishnan Unnikrishnan, Saeid Haghighatshoar, and Martin Vetterli. 2018. Unlabeled Sensing with Random Linear Measurements. *IEEE Transactions on Information Theory* 64, 5 (2018), 3237–3253.

[57] Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. Decimeter-Level Localization with a Single WiFi Access Point. In *Proc. of the 13th USENIX NSDI*. 165–178.

[58] Aditya Virmani and Muhammad Shahzad. 2017. Position and Orientation Agnostic Gesture Recognition Using WiFi. In *Proc. of the 15th ACM MobiSys*. 252–264.

[59] Fei Wang, Sanping Zhou, Stanislav Panev, Jinsong Han, and Dong Huang. 2019. Person-in-WiFi: Fine-grained Person Perception using WiFi. In *Proc. of the 33rd IEEE ICCV*. 5452–5461.

[60] Qilong Wang, Banggu Wu, Pengfei Zhu, Peihua Li, Wangmeng Zuo, and Qinghua Hu. 2020. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In *Proc. of the 33rd IEEE/CVF CVPR*. 11534–11542.

[61] Wei Wang, Alex X. Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and Modeling of WiFi Signal based Human Activity Recognition. In *Proc. of the 21st ACM MobiCom*. 65–76.

[62] Xuyu Wang, Chao Yang, and Shiwen Mao. 2017. PhaseBeat: Exploiting CSI Phase Data for Vital Sign Monitoring with Commodity WiFi Devices. In *Proc. of the 37th IEEE ICDCS*. 1230–1239.

[63] Yichao Wang, Yili Ren, Yingying Chen, and Jie Yang. 2022. Wi-Mesh: A WiFi Vision-based Approach for 3D Human Mesh Construction. In *Proc. of the 20th ACM SenSys*. 362–376.

[64] Yichao Wang, Yili Ren, and Jie Yang. 2024. Multi-Subject 3D Human Mesh Construction Using Commodity WiFi. *Proc. of the ACM IMWUT* 8, 1 (2024), 1–25.

[65] Chenhao Wu, Xuan Huang, Jun Huang, and Guoliang Xing. 2023. Enabling Ubiquitous Wi-Fi Sensing with Beamforming Reports. In *Proc. of 37th ACM SIGCOMM*. 20–32.

[66] Yaxiong Xie, Zhenjiang Li, and Mo Li. 2015. Precise Power Delay Profiling with Commodity Wi-Fi. In *Proc. of the 21st ACM MobiCom*. 53–64.

[67] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. mD-Track: Leveraging Multi-Dimensionality for Passive Indoor Wi-Fi Tracking. In *Proc. of the 25th ACM MobiCom*. 8:1–16.

[68] Jie Xiong and Kyle Jamieson. 2013. ArrayTrack: A Fine-Grained Indoor Location System. In *Proc. of the 10th USENIX NSDI*. 71–84.

[69] Jie Xiong, Karthikeyan Sundaresan, and Kyle Jamieson. 2015. ToneTrack: Leveraging Frequency-Agile Radios for Time-Based Indoor Wireless Localization. In *Proc. of the 21st ACM MobiCom*. 537–549.

[70] Ruitao Xu, Gaotao Shi, Jun Luo, Zenghua Zhao, and Yantai Shu. 2011. MuZi: Multi-Channel ZigBee Networks for Avoiding WiFi Interference. In *Proc. of the 4th IEEE/ACM CPSCOM*. 323–329.

[71] Moustafa Youssef and Ashok Agrawala. 2005. The Horus WLAN Location Determination System. In *Proc. of the 3rd ACM MobiSys*. 205–218.

[72] Youwei Zeng, Dan Wu, Jie Xiong, Jinyi Liu, Zhaopeng Liu, and Daqing Zhang. 2020. MultiSense: Enabling Multi-Person Respiration Sensing with Commodity WiFi. In *Proc. of the 22nd UbiComp*. 102:1–29.

[73] Chi Zhang, Feng Li, Jun Luo, and Ying He. 2014. iLocScan: Harnessing Multipath for Simultaneous Indoor Source Localization and Space Scanning. In *Proc. of the 12th ACM SenSys*. 91—104.

[74] Tianyue Zheng, Zhe Chen, Shujie Zhang, Chao Cai, and Jun Luo. 2021. MoRe-Fi: Motion-robust and Fine-grained Respiration Monitoring via Deep-Learning UWB Radar. In *Proc. of the 19th ACM SenSys*. 111–124.

[75] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2019. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi. In *Proc. of the 17th ACM MobiSys*. 313–325.

[76] Zheng Yang and Chenshu Wu and Yunhao Liu. 2012. Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention. In *Proc. of the 18th ACM MobiCom*. 269–280.